

1. При комплексном обследовании нескольких тысяч человек по измерявшимся показателям (включая пульс, давление, ЭКГ и т.д.) оценивался риск возникновения сердечного заболевания. Ста пациентам с самым высоким риском была предложена оздоровительная программа, включающая диету, упражнения и приём профилактических препаратов. Через несколько месяцев после окончания программы пациенты снова прошли диспансеризацию; средний оцениваемый риск возникновения сердечного заболевания существенно уменьшился.

1 / 1 point

Что можно сказать об эффективности оздоровительной программы?

- ☐ Программа явно эффективна — риск уменьшился, значит, пациенты стали здоровее!
- ☒ Данных недостаточно: поскольку были выбраны пациенты с наибольшим риском, измеренный эффект может объясняться регрессией к среднему. Для оценки эффективности программы нужно использовать контрольную группу пациентов с таким же высоким риском, для которых программа не проводилась, и сравнить изменения в двух группах.

✓ Correct

2. Из 15 клиентов банка, которым менеджер предложил подключить автоплатёж, четверо согласились. Подключение услуги — бинарный признак, который можно описать распределением Бернулли. Запишите функцию правдоподобия $L(X^n, p)$ для такой выборки и посчитайте её значение, если истинная вероятность подключения услуги $p = 0.2$. Запишите ответ с пятью знаками после десятичной точки.

0 / 1 point

2

! Incorrect

Значение функции правдоподобия равно вероятности выборки, а вероятность не может быть больше единицы.

3. По выборке из предыдущей задачи найдите оценку максимального правдоподобия для параметра p . Запишите ответ с тремя знаками после десятичной точки.

0 / 1 point

3

! Incorrect

4. Выберите верные утверждения о регрессии, получаемой методом наименьших квадратов.

1 / 1 point

- ☒ Если шум описывается нормальным распределением с нулевым средним и постоянной дисперсией, метод наименьших квадратов даёт оценку максимального правдоподобия.

✓ Correct

- ☒ Получаемая оценка приближает условное по x матожидание отклика y .

✓ Correct

- ☐ Получаемая оценка приближает условную по x медиану отклика y .

- ☐ Если шум описывается лапласовским распределением с нулевым средним и постоянной дисперсией, метод наименьших квадратов даёт оценку максимального правдоподобия.

5. Какие из приведённых ниже средств помогают от переобучения линейных моделей?

1 / 1 point

- ☐ Усложнение модели.

- ☒ Регуляризаторы.

✓ Correct

- ☒ Увеличение количества данных.

✓ Correct

- ☒ Упрощение модели.

✓ Correct

- ☐ Использование средней абсолютной ошибки вместо среднеквадратичной.

6. Выберите верные утверждения о регрессии, получаемой методом наименьших квадратов с L_1/L_2 регуляризаторами.

1 / 1 point

- ☒ Из-за регуляризаторов оценки коэффициентов модели получаются смещённые.

✓ Correct

- ☐ Регуляризация подходит только для линейной регрессии, к сожалению, в задачах остальных типов её использовать нельзя.
- ☒ Регуляризаторы позволяют бороться с переобучением модели.

✓ Correct

- ☐ Оба типа регуляризаторов уменьшают по модулю веса признаков, но L_2 некоторые ещё и обнуляет.
- ☒ Константное слагаемое не должно входить в регуляризатор.

✓ Correct

Потому что штраф на него никак не избавляет от переобучения, но при этом ухудшает качество модели.

- ☐ Регуляризаторы увеличивают дисперсию оценок коэффициентов регрессионной модели.
- ☒ При использовании L_2 -регуляризатора в МНК решение можно найти аналитически, а с L_1 -регуляризатором — только численно.

✓ Correct

7. Выберите верные утверждения о логистической регрессии.

1 / 1 point

- ☐ Логистическая регрессия позволяет построить линейную модель вероятности $P(y = 1 | X)$
- ☐ Функция, минимизируемая при настройке параметров модели, невыпуклая, у неё много локальных экстремумов, поэтому задача всегда имеет много решений.
- ☒ При настройке моделей могут возникать проблемы, если классы линейно разделимы в пространстве признаков.

✓ Correct

В таком случае $\|w\| \rightarrow \infty$

- ☒ Оценка параметров модели делается минимизацией кросс-энтропии.

✓ Correct

Минимизация кросс-энтропии — то же самое, что максимизация правдоподобия.